

Домашнее задание №1

Новик Никита, АСОБОИ-1

20 сентября 2017 г.

Задание 1

Пункт 1.1 Условие

$f : (0, +\infty) \rightarrow R$, обратима

$X - \text{CB}$

Доказать, что $\forall t > 0 : P(X > t) \leq f(t) \Rightarrow \forall \delta > 0 : P(X \leq f^{-1}(\delta)) \geq 1 - \delta$

Пункт 1.2 Решение

Рассмотрим следующее выражение:

$$0 \leq P(X > t) \leq f(t). \quad (1)$$

Из этого следует:

$$P(X \leq t) \geq 1 - f(t). \quad (2)$$

Так функция $f(t)$ обратима, то δ можно представить в виде $\delta = f(t)$. Из этого получаем:

$$P(X \leq f^{-1}(\delta)) \geq 1 - \delta \quad (3)$$

Задание 2

Для того, чтобы доказать что в классе h_p найдется классификатор, совпадающий с h_S , достаточно привести алгоритм построения многочлена для h_p по h_S .

Вначале выберем все такие x_{ik} , на которых $h_S(x_{ik}) = 1$. Затем построим многочлен

$$-(x - x_{i1})^2(x - x_{i2})^2 \dots (x - x_{ik})^2.$$

Этот многочлен принимает только отрицательные значения на всем $X = R$, кроме точек x_{ik} , где достигаются нулевые значения. Таким образом, мы построили такой многочлен для h_p , чтобы классификаторы совпадали.

Так как в классе h_p найдется классификатор, совпадающий с h_S , то при применении ERM-парадигмы в классе h_p может произойти переобучение. А это значит, что применение ERM-парадигмы может приводить к плохим результатам.

Задание 3

Пункт 3.1

Пусть алгоритм А выбирает наименьший прямоугольник, содержащий все точки положительного класса. Построим с его помощью по произвольной тестовой выборке S гипотезу h^* . Так как выполняется предположение о реализуемости и эмпирический риск $L_S(h^*) = 0$, то отсюда следует, что А является реализацией ERM-алгоритма.

Пункт 3.2 Смотреть в third_task.ipynb

Пункт 3.3 Смотреть в third_task.ipynb

Пункт 3.4 Смотреть в third_task.ipynb

$$10\% \rightarrow 42, 1\% \rightarrow 430, 0.1\% \rightarrow 3100$$

Пункт 3.5

При увеличении площади X необходимый размер выборки уменьшается.

При пропорциональном увеличении площади Q и X останется примерно таким же.

При увеличении размерности X останется примерно таким же.

Результат должен зависеть от D , так как качество тренировочной выборки S зависит от D , вероятность нерепрезентативной выборки δ зависит от S , а m зависит δ .